

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
7 April 2005 (07.04.2005)

PCT

(10) International Publication Number
WO 2005/031654 A1

(51) International Patent Classification⁷: **G06T 15/70,**
G10L 21/06

(21) International Application Number:
PCT/IB2004/051903

(22) International Filing Date:
28 September 2004 (28.09.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/507,098 30 September 2003 (30.09.2003) US

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS, N.V.** [NL/NL];
Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DIMITROVA,**

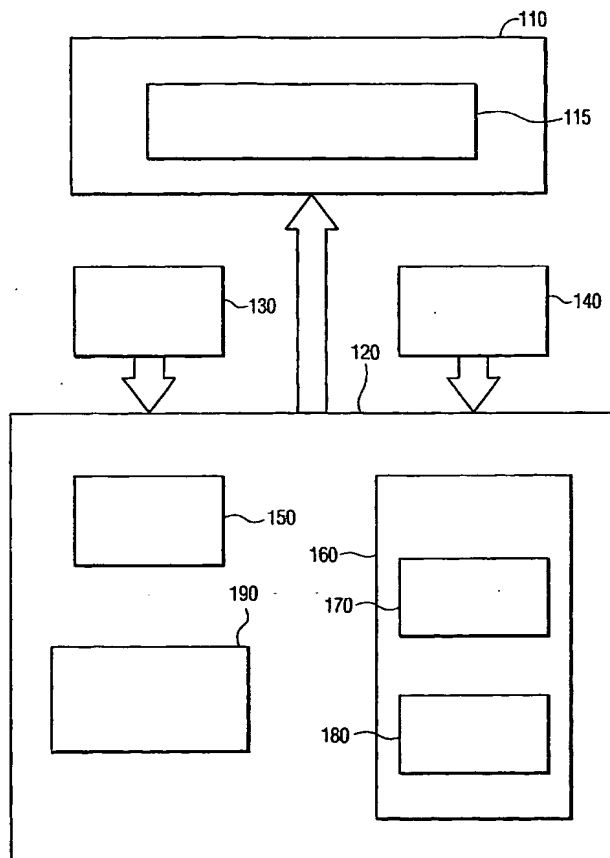
Nevenka [MK/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **MILLER, Andrew** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **LI, Dongge** [CN/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(74) Common Representative: **KONINKLIJKE PHILIPS ELECTRONICS, N.V.**; c/o Waxler, Aaron, P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR AUDIO-VISUAL CONTENT SYNTHESIS



(57) Abstract: A system and method is provided for synthesizing audio-visual content in a video image processor. A content synthesis application processor extracts audio features and video features from audio-visual input signals that represent a speaker who is speaking. The processor uses the extracted visual features to create a computer generated animated version of the face of the speaker. The processor synchronizes facial movements of the animated version of the face of the speaker with a plurality of audio logical units such as phonemes that represent the speaker's speech. In this manner the processor synthesizes an audio-visual representation of the speaker's face that is properly synchronized with the speaker's speech.



(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ,

TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.